



# Multi-objective calibration of the land surface scheme TERRA/LM using LITFASS-2003 data

K.-P. Johnsen, H.-T. Mengelkamp, S. Huneke

## ► To cite this version:

K.-P. Johnsen, H.-T. Mengelkamp, S. Huneke. Multi-objective calibration of the land surface scheme TERRA/LM using LITFASS-2003 data. *Hydrology and Earth System Sciences Discussions*, 2005, 9 (6), pp.586-596. hal-00304871

**HAL Id: hal-00304871**

**<https://hal.science/hal-00304871>**

Submitted on 31 Dec 2005

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Multi-objective calibration of the land surface scheme TERRA/LM using LITFASS-2003 data

K.-P. Johnsen, H.-T. Mengelkamp and S. Huneke

GKSS Research Centre, Institute for Coastal Research, D-21502 Geesthacht, Germany

E-mail for corresponding author: Johnsen@gkss.de

### Abstract

The turbulent sensible and latent heat fluxes simulated in the operational weather forecast model LM have been checked with data from the field experiment LITFASS 2003 (Lindenberg Inhomogeneous Terrain – Fluxes between Atmosphere and Surface: a Long-term Study) using both single site measurements and grid box aggregated fluxes. SCE-UA (single objective) and MOSCEM-UA (multi-objective) approaches were applied to calibrate the land-surface scheme TERRA/LM for 11 single sites and for the aggregated fluxes. A large variation is seen among the parameter sets found by calibration but no typical classification according to vegetation type is obvious. This is attributed to the calibrated parameter sets correcting for model deficiencies and data errors rather than describing the physical characteristics of the measurement site. The measured fluxes were combined into a time series of aggregated fluxes by the tile method. Calibration of TERRA/LM with respect to the averaged fluxes resulted in a range of parameter sets which all simulated the area-averaged fluxes in much better agreement with the observed fluxes than the standard parameter set of the operational model. A modified Nash-Sutcliffe measure as a coincidence criterion fell from 0.3 to a range between 0.15 and 0.28 for the latent heat flux and from 0.43 to between 0.26 and 0.36 for the sensible heat flux when the calibrated parameter sets were used instead of the standard parameters.

**Keywords:** sensible heat, latent heat, LITFASS, TERRA/LM, objective calibration

### Introduction

Numerical weather prediction (NWP) models contain land surface schemes (LSS) to describe the heat and water fluxes at the land surface atmosphere interface and within the soil. The development of these LSS has resulted in highly complex models with many parameters, some of which can be measured but many are generally unknown. The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS; Henderson-Sellers *et al.*, 1995) has shown that different LSS give significantly different surface fluxes because of different parameter sets and different model structures.

The turbulent exchange of energy and water vapour between the land surface and the atmosphere is influenced by the complex heterogeneity of the land surface due to different scales of variability of the vegetation, the surface type and the orography as well as the interaction with temporally and spatially varying meteorological variables. Spatial heterogeneous processes of scale smaller than the

grid size of the NWP model have to be parameterised. Over the last 20 years, different approaches to the description of sub-grid-scale processes in atmospheric and hydrological models have been developed and two have been considered:

1. The assumption that sub-grid-scale processes influence larger scale processes significantly and that non-linearities of the sub-scale processes do not allow a simple scaling (Avissar and Pielke, 1989), and
2. the conclusion from measurements that land surface models are mainly independent of the scale and approaches with effective parameters are suitable (Wood, 1995).

The application of multi-criteria optimisation methods to adjusting the model parameters of a LSS have been shown to lead to more realistic simulations of the surface energy fluxes for different types of land cover (Gupta *et al.*, 1999; Houser *et al.*, 2001; Vrugt *et al.*, 2003).

In this study, single-criterion and multi-criteria methods have been applied to find parameter sets for the LSS TERRA/LM (the land surface scheme in the operational weather forecast model LM of the German Weather Service (DWD)) which appropriately describe the area-averaged surface energy fluxes over a typical mid-European heterogeneous vegetation cover and at single sites of the LITFASS (Lindenberg Inhomogeneous Terrain – Fluxes between Atmosphere and Surface: a Long-term Study) 2003 research experiment (Beyrich *et al.*, 2004).

## Data and models

## THE LITEASS 2003 EXPERIMENT

The LITFASS 2003 research experiment is central to the German Climate Research Program (DEKLIM) project EVA-GRIPS (EVAporation at the GRId and Pixel Scale), the main aim of which is the determination of evaporation over a typical mid-European heterogeneous land surface at the meso-scale, i.e. the grid-scale of a regional NWP model (roughly 5 to 10 km grid length) (Mengelkamp *et al.*, 2004) and its description within LSS. The LITFASS 2003

experiment was located in a  $21 \times 21 \text{ km}^2$  area close to the Meteorological Observatory Lindenberg (MOL; Germany) of the DWD, close to the city of Berlin, between May 19 and June 17, 2003 (Fig. 1). The spatial distribution of the land use classes at the LITFASS 2003 site is shown in Fig. 2 and the fraction of each land use class in Table 1. The area covers a square of  $3 \times 3$  grid cells of the Lokalmmodell.

Ground-based eddy correlation instruments, an infrared camera for surface photography on board a Tornado RECCE aircraft of the German air force (Bange *et al.*, 2004), a helicopter-borne turbulence probe HELIPOD and infrared data from the AVHRR sensors of NOAA satellites were used to estimate the energy and water fluxes at different scales. The land surface and the boundary layer of a typical, natural and heterogeneous landscape of northern Europe were investigated at a total of 14 micrometeorological stations. The radiation balance, soil variables (temperature, humidity, heat flux) and the air temperature, pressure, humidity, wind speed, wind direction and precipitation were measured together with the surface fluxes. The stations were erected above nine agricultural sites (barley (2 stations); rape (3); triticale, a hybrid of wheat and rye (2); and maize (2)), two grassland sites, one pine forest site (canopy height 14 m),

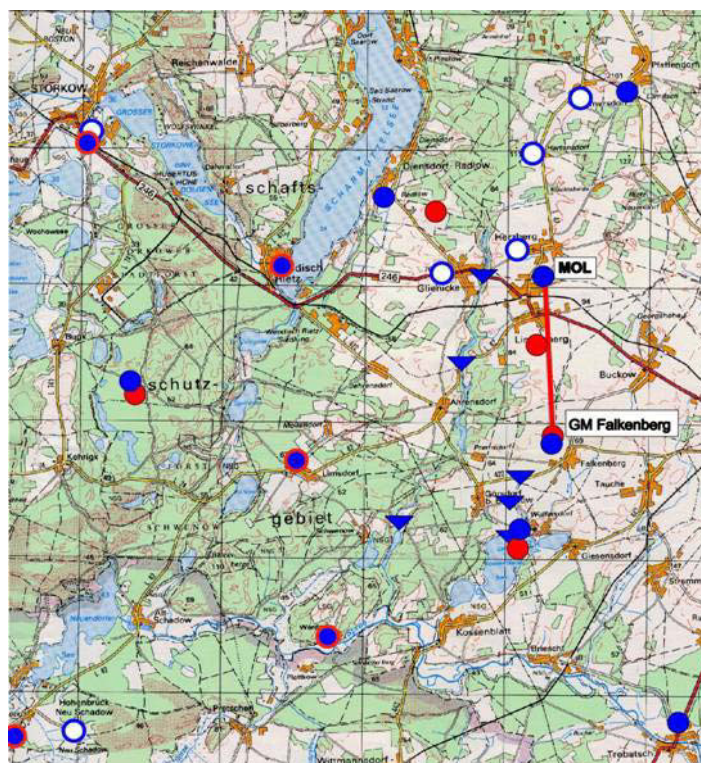


Fig. 1. LITFASS 2003 area with the measurement stations. Surface based measurements of the turbulent fluxes were taken at 13 sites (red dots), ground-based remote sensing systems were operated at three sites (yellow dots), the blue symbols mark the position of rain gauges of a regional precipitation network (surrounded by a red ring where global radiation is also measured), and the red lines indicate the long-distance scintillometer paths.

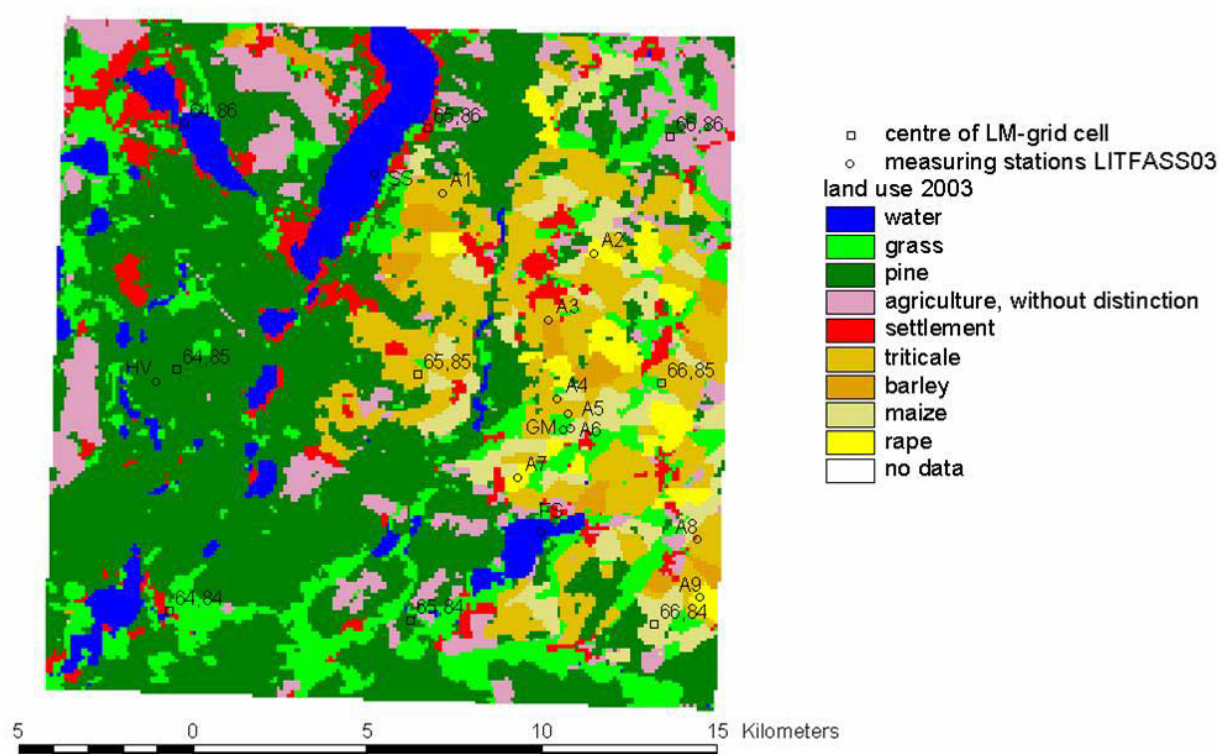


Fig. 2. Land use classes of the LITFASS area, 2003.

two lake sites, and at two levels of a 100 m tower (Beyrich *et al.*, 2004b). The net short- and long-wave radiation balance, the specific humidity and air temperature were measured at heights between 1.9 and 2.3 m with standard instrumentation. The turbulent fluxes of momentum, sensible and latent heat were measured between 2.8 and 3.5 m above ground. To minimise differences in the estimates of the turbulent surface fluxes from the eddy-covariance measurements caused by different sensor and system characteristics, only two types of ultrasonic anemometer-thermometers and fast response hygrometers were used, namely sonics CSAT-3 (Campbell Sci. Ltd.) and

USA-1 (METEK GmbH), and hygrometers KH-20 (Campbell Sci. Ltd.) and LI7500 (LiCor Inc.). A laboratory calibration procedure for the fast response hygrometers was set up and tested at MOL. Prior to and after LITFASS-2003, all KH-20 and LI7500 instruments from the different groups were calibrated with this unified procedure to ensure comparability of the computed fluxes with respect to data treatment and correction algorithms and high quality turbulent flux measurements (Foken and Oncley, 1995; Foken and Wichura, 1996; Foken *et al.*, 2004). The major components of this quality control system comprise conversion of the high frequency raw data into meteorological units, a head correction for the METEK USA-1 sonic anemometers (HC1) obtained from wind tunnel studies by the manufacturer, a correction procedure for crosswind effects and the coordinate transformation (Wilczak *et al.*, 2001). The fast temperature and humidity data were corrected for noise temperature fluctuations, density fluctuations and oxygen sensitivity. All the turbulence measurements were tested for electrical and physical plausibility using consistency limits for each measured variable. Finally, the fluxes were checked for stationarity and for integral turbulence characteristics.

The focus of this paper is on the calibration of a land-surface scheme, which inherently assumes closure of the surface energy balance, although it is well known that

Table 1. Partitioning of land-use classes in the LITFASS 2003 area

Land-use class	Percent
Water	7
Forest	43
Grass	13
Triticale	17
Barley	3
Rape	4
Maize	8
Settlement	5



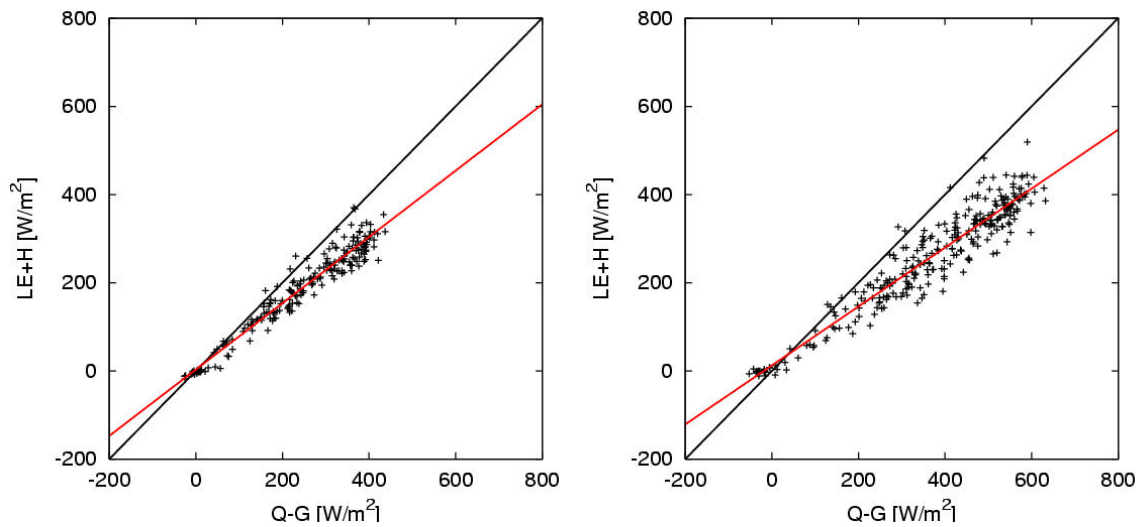


Fig. 3. Comparison of the sum of the sensible and latent heat flux to the difference between net radiation and soil heat flux for the station with the lowest energy balance gap (left) and the station with the largest gap (right).

measurements seldom show such a closure. In Fig. 3, the sum of the sensible and latent heat fluxes is compared with the difference of net radiation and soil heat flux from 30-min measured averages at a grassland site (GM) which had the smallest gap in the energy balance closure and at a rape site (A9) for which the gap in closure was the largest of all stations. The RMSE of all stations varies between  $70 \text{ Wm}^{-2}$  and  $130 \text{ Wm}^{-2}$  and the bias was between  $57 \text{ Wm}^{-2}$  and  $107 \text{ Wm}^{-2}$ . Although no comment is made on the year-long discussion on the source of the gap in closure of the energy balance, the present analysis of the parameter calibration for a land-surface scheme has been based on observations showing a balance closure of only about 70% in the worst case for 30-min averages.

#### THE LAND SURFACE SCHEME LM/TERRA

The LSS LM/TERRA is incorporated in the non-hydrostatic numerical weather prediction model LM (Lokalmodell) of the DWD (Schrodin and Heise, 2001); a stand-alone version of LM/TERRA, extracted from its host model, is driven by downward short- and long-wave radiation, air temperature and humidity and wind speed as well as surface pressure and precipitation. LM/TERRA calculates the sensible and latent heat fluxes and the momentum flux. The transfer coefficients for momentum and heat are derived from Monin-Obukhov similarity theory with the iterative method of Louis (1979). The transfer coefficients for momentum and heat are identical in form as:  $c = u \cdot \left( \frac{\kappa}{\ln(h/z_0)} \right) \cdot f$  with  $u$  the wind speed and  $f$  the appropriate stability function;  $h$  is the height above ground level,  $\kappa$  the von Karman constant and  $z_0$  the roughness length.

At the upper soil boundary, the soil heat flux is calculated as the residuum of the radiation balance and the turbulent sensible and latent heat fluxes as is common to land-surface schemes as they inherently close the surface energy balance. However, as was stated above, the closure is not *a priori* given for the measurements. Because calibration has been effected with respect to the turbulent heat fluxes, there might well be larger differences between the simulated and measured soil heat flux.

The heat and water transport in the soil is calculated for six layers, with boundaries at 1, 4, 10, 22, 46 and 94 cm depth. Numerical experiments have shown that at least four active layers should be used for numerical weather prediction models (Schrodin and Heise, 2001). The temperature is determined by solving a multi-layer version of the heat conduction equation:  $\frac{\partial T_{so}}{\partial t} = \frac{1}{c_s} \cdot \frac{\partial}{\partial z} \left( \lambda \frac{\partial T_{so}}{\partial z} \right)$ . Here  $T_{so}$  is the soil temperature,  $t$  the time,  $c_s$  the soil heat capacity,  $\lambda$  the heat conductivity and  $z$  depth. At the lower boundary, a constant (climate value) is used for the temperature. Soil moisture diffusion is calculated using the Richards equation with no drainage as the lower boundary condition. To determine the evaporation from bare soil and the transpiration by plants, the model uses the BATS scheme of Dickinson (1986) and water and energy transport between the soil and the plant canopy are neglected. This leads to a Monteith combination formula for the computation of plant transpiration. If the potential evaporation is less than zero at the surface, the total plant transpiration  $T$  is:

$$T = v_R \cdot (1.0 - f_i) \cdot E_{pot}(T_{sf}) \cdot r_a \cdot (r_a + r_f)^{-1}$$

where  $v_R$  is the vegetation ratio,  $f_i$  the fractional area of the interception storage,  $E_{pot}$  the potential evaporation,  $T_{srf}$  the surface temperature and  $r_a$  and  $r_f$  are the atmospheric and foliage resistances, respectively. The stomata resistance  $R_{stom}$  is derived according to:

$$R_{stom}^{-1} = (R_{stom}^{\max})^{-1} + \left( (R_{stom}^{\min})^{-1} - (R_{stom}^{\max})^{-1} \right) F_{Rad} \cdot F_{Wat} \cdot F_{Temp} \cdot F_{Sat}$$

with the maximum stomatal resistance  $R_{stom}^{\max} = 5000 \frac{s}{m}$ ,  $R_{stom}^{\min}$  the variable minimum stomatal resistance. The radiation function  $F_{Rad}$  describes the influence of the photosynthetically active radiation and  $F_{Wat}$ ,  $F_{Temp}$  and  $F_{Sat}$  take into account the effects of water stress, water deficit and air temperature on the surface resistance, respectively (Noilhan and Planton, 1989). All F-Functions of the stomatal resistance are limited between zero and one.

Water after rain is stored in an interception storage of maximum capacity  $W_{i,max}$  depending on the fractional area of the vegetation  $v_R$  and a minimum value for the maximum interception water content  $W_{i,0}$ :  $W_{i,max} = W_{i,0} (1.0 + 5.0 v_R)$ . If the water content of the interception store is greater than zero, part of the water intercepted will percolate with time constant  $\tau_{perc}$  to the uppermost soil layer. The maximum infiltration rate  $I_{max}$ , given by a simplified Holtan equation, is zero if the surface temperature is lower than or equal to 273.15 K:

$$I_{max} = f \left( \max(0.5, v_R) \cdot I_{k1} \frac{w_{pv} - w_1}{w_{pv}} + I_{k2} \right)$$

where  $w_{pv}$  the pore volume and  $w_1$  the water content of the uppermost soil layer.  $I_{k1}$  equals 0.002 and  $I_{k2}$  depends on the soil type. Most variables in the soil model (heat capacity, water storage capacity, etc.) depend strongly on soil texture. Out of the five different soil texture classes (sand, sandy loam, loam, loamy clay, and clay), sandy loam was used for the LITFASS area. For lakes, no soil model was applied but the water temperature was prescribed according to measurements. For further details see the Documentation of the EM/DM system (DWD, 1995).

#### THE MULTI-OBJECTIVE CALIBRATION ALGORITHM

While manual or automated calibration is common in hydrology (Gupta *et al.*, 1998) less effort is devoted to determining the parameter uncertainties in land surface modelling. The recently developed multi-criteria calibration algorithm (Bastidas *et al.*, 1999, 2001) allows the model parameter space to be restricted to a set of globally optimal and equally qualified parameter sets. The algorithm, called Multi-Objective Shuffled Complex Evolution Metropolis algorithm (MOSCEM-UA), was developed at the Universities of Amsterdam and Arizona (Vrugt *et al.*, 2003;

Gupta *et al.*, 1999). It is based on the complex shuffling applied in the SCE-UA algorithm developed by Duan (1994) and the probabilistic covariance-annealing search procedure of the SCEM-UA algorithm (Vrugt *et al.*, 2003). It allows optimisation problems to be solved with more than one completely independent objective function. As independent objective functions  $OF$ , a modified Nash-Sutcliffe measure (optimum at 0) was applied to the differences between the measured ( $Q_{obs}$ ) and calculated heat fluxes ( $Q_{sim}$ ):

$$OF = \frac{\sum_{i=1}^N (Q_{obs} - Q_{sim})^2}{\sum_{i=1}^N (Q_{obs} - \overline{Q_{obs}})^2}$$

Gupta *et al.* (1998) list some other possible objective functions but, in over 20 years of investigation, it has not proved possible to demonstrate clearly that one particular objective function is better suited for calibration than another (Gupta *et al.*, 1998). Here, two objective functions were used for the latent and sensible heat fluxes, respectively. The algorithm determines the set of pareto optimal objective function vectors with pareto rank 1 (Goldberg, 1989). A vector  $\mathbf{x}$  of objective functions is said to dominate (i.e. has a lower rank than) another objective function vector  $\mathbf{y}$  if an  $i$  exists with  $x_i < y_i$  for all  $x_i$  and for all  $i$ ,  $x_i \leq y_i$ . Here  $x_i$  and  $y_i$  are the elements of the vectors  $\mathbf{x}$  respectively  $\mathbf{y}$ . The pareto-optimal objective function vectors of rank 1 are non-dominated. The MOSCEM-UA algorithm takes an initial population of points randomly spread in the feasible parameter space. For each individual of the population, the multi-objective vector is computed and the population is ranked and sorted using a method similar to the fitness assignment concept of Zitzler and Thiele (1999). The population (here with 500 individuals) is partitioned into  $q$  complexes (here  $q=10$ ) and in each complex parallel sequences of individuals are launched and new candidate points are derived from the structure of the sequences and the associated complex. A Metropolis-Hasting algorithm will decide if the candidate points are accepted and, in the case of acceptance, will replace the worst member of the current complex. After a prescribed number,  $k$ , of iterations (here  $k=10\,000$ ) the complexes are combined and new complexes are formed through a process of shuffling (Vrugt *et al.*, 2003). The process converges to a so-called 'Pareto set'. There is no unique minimising point in the parameter space because of deficiencies in model physics and errors in the observations. A characteristic of the set of solutions in the Pareto set is the fact that each parameter set is better than the others for at least one of the objective functions, but no one is better than another for all objectives. The Pareto set forms a curve in the space of the objective functions with the length of the curve characterising the precision of the

calibration while the position characterises the accuracy of the simulation.

## Results

The main purpose of this study is the verification of the grid-averaged turbulent sensible and latent heat fluxes in the operational weather forecast model LM using data from the field experiment LITFASS 2003. The results are presented in two steps.

The SCE-UA (single objective) and the MOSCEM-UA (multi objective) approaches are applied to calibrate the land-surface scheme TERRA/LM for 11 single sites in the area of the LITFASS 2003 experiment. The parameters selected, their prescribed numerical ranges and their respective values as given in the standard TERRA/LM are listed in Table 2. With the parameters found by calibration (Table 3) TERRA/LM is validated using an alternative data set from the same sites, using the procedure described earlier.

In a second step, TERRA/LM is calibrated to the LITFASS 2003 grid square averaged fluxes. Time series of area representative fluxes are built from the point measurements by weighting them according the occurrence of the respective land use class (Table 1), the so-called tile approach. These time series are then used to calibrate TERRA/LM to find aggregate parameters appropriate to describe the area averaged fluxes. Fluxes derived from the measurements by applying the tile approach are compared to those simulated with the standard TERRA/LM parameter set and with the aggregate parameters.

### SITE SPECIFIC CALIBRATION AND VALIDATION

In this study, the single-criterion SCE-UA algorithm (Duan, 1994) and the multi-criteria algorithm MOSCEM-UA have been applied consecutively. SCE-UA is used to calibrate a total of 12 parameters as listed in Table 3 while, in a second step, MOSCEM-UA is used to calibrate a subset of 6

Table 2. Optimal parameter sets of the TERRA/LM model for the different land surface sites:  $z_0$  is the roughness length,  $\alpha_{sw}$  the short-wave albedo,  $R_{s,min}$  the minimum stomatal resistance, veg is the vegetation ratio, LAI the leaf area index,  $c_s$  the soil heat capacity, and  $v_{FC}$  the field capacity.

Parameter	Range	Standard LM	barley	grass NV2	grass NV4	maize	maize UBT	rape	rape WAU	triticale UBT	triticale WAU	Forest	Area averaged
$z_0$	0.01..0.8 m	0.03	70.149	0.149	0.010	0.010	0.0255	0.224	0.790	0.151	0.192	0.245	0.3467
$\alpha_{sw}$	0.05..0.35	0.15	0.334	0.314	0.348	0.350	0.340	0.337	0.348	0.282	0.324	0.132	0.210
$R_{s,min}$	20..500 s/m	60	60.01	86.89	60.01	30.01	192.9	20.82	20.31	24.01	24.58	38.02	92.92
Veg	0..1	1.0	0.0171	0.615	0.536	0.987	0.621	0.935	0.998	0.664	0.950	0.568	0.657
LAI	1..6	3.964	4.845	3.792	5.130	4.751	3.664	4.814	5.948	4.487	2.539	3.991	4.061
$c_s$	500..5000 Jkg <sup>-3</sup> K <sup>-1</sup>	1280	1120	3170	517	658	1810	3178	632	702	535	1460	2180
$v_{FC}$	0.1..0.9	0.196	0.206	0.105	0.128	0.113	0.135	0.359	0.211	0.182	0.116	0.838	0.139

Table 3. Optimal parameter sets of the TERRA/LM model for the different land surface sites from the single-objective calibration:  $k_{root}$  is the root depth,  $\tau_{perc}$  the time constant for percolation of water from the interception store to the uppermost soil layer, and  $W_{i,0}$  is the parameter for the maximum interception water content.  $F_{rad}$  is the radiation function (or fraction of the photosynthetic active radiation) and  $V_p$  is the pore volume.

Parameter	Range	barley	grass NV2	grass NV4	maize	maize UBT	rape	rape WAU	triticale UBT	triticale WAU	Forest	Area averaged
$k_{root}$	0.01...1.00 m	0.455	0.413	0.578	0.477	0.386	0.334	0.454	0.352	0.518	0.273	0.500
$\tau_{perc}$	1200...5 000 000 s	31290	226000	10700	3369	7140	54051	1216	5709	11832	9157	5402
$W_{i,0}$	0.0005...0.05 kgm <sup>-2</sup>	0.00054	0.0012	0.032	0.016	0.0036	0.027	0.021	0.0097	0.028	0.026	0.033
$F_{rad}$	0.4...0.5	0.430	0.442	0.466	0.418	0.429	0.445	0.499	0.427	0.450	0.463	0.449
2	0...1	0.925	0.303	0.828	0.514	0.566	0.370	0.2114	0.740	0.589	0.718	0.525

parameters (Table 2). The target single objective function for the SCE-UA calibration is the sum of the squared Nash-Sutcliffe measures of latent and sensible heat. The target multi (two) objective functions for the MOSCEM-UA calibration are the Nash-Sutcliffe measures of the latent and sensible heat, respectively. This procedure was adopted because of the enormous computer resources required for the multi-objective MOSCEM-UA calibration for as many as 12 parameters in a reasonable time. The parameters listed in Tables 2 and 3 were subject to the single objective calibration. Then the parameters in Table 3 were fixed and only the parameters in Table 2 were calibrated by the multi-objective approach. The parameters selected for the final multi-objective calibration (Table 2) should characterise the surface (roughness length and albedo), the vegetation (minimum stomatal resistance and leaf area index) and the soil (soil heat capacity and field capacity). Their ranges are prescribed according to measurements taken during the LITFASS 2003 experiment (Beyrich *et al.*, 2004). For all parameters without measured ranges, the boundaries were chosen so that the physical meaning was retained and the maximum ranges were identical for each land-use class.

All parameters are assumed to be time invariant over the measurement period between May 19 and the June 17, 2003, although this is the main growth season. The optimal parameters are site specific rather than land-use class specific, firstly because the algorithm is able to find the global optimum despite the high complexity of the LSS combined with measurement uncertainties and secondly because, within the parameter spaces, many quite similar local optima exist with completely different parameter sets. Nevertheless, the parameter sets allow TERRA/LM to describe the given data sets with the optimal Nash-Sutcliffe efficiencies.

The parameters in Table 2 found by the multi-objective calibration describe the objective functions closest to the lower left corner in Fig. 4. They show quite diffuse behaviour, with no clear distinction between the different land use classes and large differences occur for sites with identical vegetation types. The roughness length for grass varies from 0.01 m to 0.15 m and for rape fields from 0.22 to 0.79. The value for the forest site is much lower than expected on physical grounds. The minimum stomatal resistance shows the largest difference between the two maize sites. Note that in the process of calibration the parameters may, to a certain extent, lose their physical meaning and so may correct for model deficiencies and data errors rather than describe the physical characteristics of the site specific soil and vegetation.

Figures 4 and 5 show the objective functions of Pareto rank 1 for all single sites for the calibration and validation periods, respectively. The calibration period was all odd dates from May 19 until June 17, 2003 and validation was all even dates from May 20 until June 16, 2003. A sufficient spin-up time was assured by doubling the data set and analysing only the second half. It was assumed that the measurements within the verification period were independent of those within the calibration period. The modified Nash-Sutcliffe measure is optimal at 0. The elements of the Pareto rank 1 objective function are qualified equally because of the two different criteria applied.

The Pareto curve for the lake station is almost a point due to the non-existence of soil and vegetation parameters. Among the various stations, the precision of the calibration (given by the length of the Pareto curve) and the accuracy of the simulation (given by the position) show some variation but, with objective functions below 0.4, the combination of model physics and parameter set describes the observed

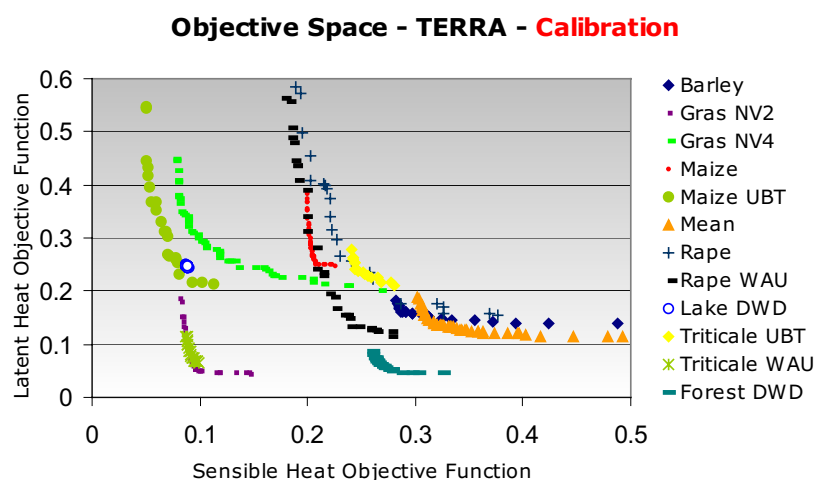


Fig. 4. Objective space for the LITFASS 2003 parameter sets of the SVAT scheme TERRA/LM with pareto rank 1 (calibration period).



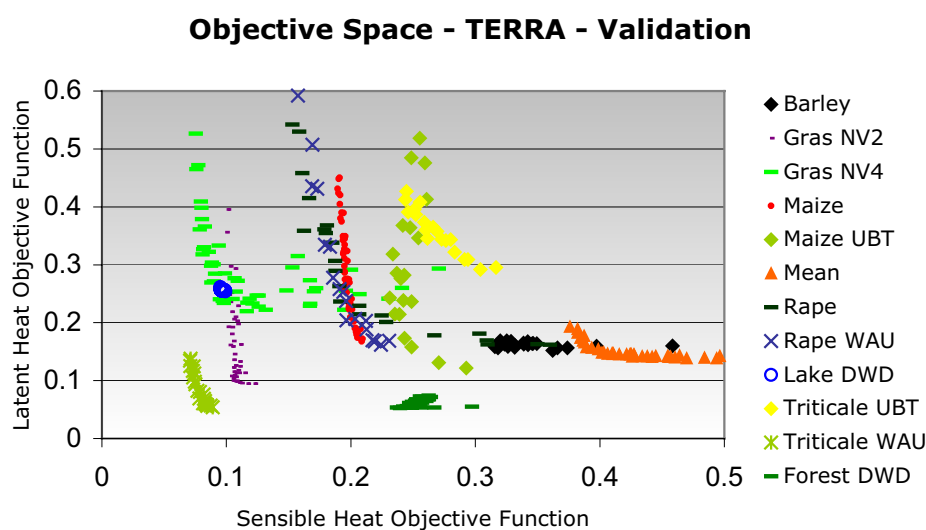


Fig. 5. Objective space for the LITFASS 2003 parameter sets of the SVAT scheme TERRA/LM with pareto rank 1 (validation period).

fluxes well and consistently for the sensible and latent heat, respectively. The sophistication of the model physics for both processes seems to be similar. The Pareto curves for the validation period show rather more scatter but are generally similar to those for the calibration period.

#### CALIBRATION AND VALIDATION OF AREA AVERAGED FLUXES

After showing that TERRA/LM with parameter sets found by calibration reproduces the sensible and latent heat flux for the single sites of the LITFASS\_2003 experiment with good coincidence with the observations, the same procedure has been applied to the area averaged fluxes to compare the calibrated model settings with those in the operational weather forecast model ('standard LM'). The LITFASS-2003 experimental area represents one grid box of the operational model.

The grid box representative time series of the turbulent fluxes are calculated from the point measurements by the so-called tile approach. The calibration process results in a parameter set representative for the whole area, the so-called 'effective parameter'. Actually there is one parameter set for each of the points in the Pareto curve (denoted 'mean' in Figs. 4 and 5). The parameters listed in Tables 2 and 3 as 'grid averaged' belong to that point in the Pareto curve closest to the coordinate origin. All parameters are well inside the range of the single site parameters. Compared to the parameters of the standard TERRA/LM version the largest differences occur for the roughness length and the minimum stomata resistance. The length and position of the Pareto curves for the area averaged fluxes ('mean', Fig. 4 and 5) indicate that the precision of the calibration and the

accuracy of the simulation are of similar order as for the single sites. This is particularly true for the latent heat flux with objective functions below 0.2 while the highest objective functions are found for the sensible heat flux between 0.3 and 0.5.

Those grid averaged fluxes deduced from the observations, those simulated with the 'standard TERRA/LM' parameter set and those simulated with all parameter sets of Pareto rank 1 are shown in Figs. 6 (sensible heat flux) and 7 (latent heat flux) for the verification days in the period May 20 to June 17, 2003. Those lines for the Pareto set 1 fluxes essentially show the range of fluxes simulation with each parameter set in Pareto rank 1. The standard TERRA/LM tends to underestimate the sensible heat flux but overestimates the latent heat flux. Obviously the standard parameter set does not allow TERRA/LM an appropriate partition of the incoming solar radiation. With the calibrated parameter sets, the simulated fluxes are in excellent agreement with the observations. For the case of the standard parameter set, the modified Nash-Sutcliffe efficiency is 0.30 for the latent heat flux and 0.38 for the sensible heat flux. With the optimised parameter sets the efficiency is reduced to a range of 0.15 and 0.28 for the latent heat flux and between 0.26 and 0.36 for the sensible heat flux for all members of the Pareto set.

## Conclusions

Objective calibration procedures have been applied to derive appropriate parameter sets for the land surface scheme TERRA/LM to simulate, realistically, the sensible and latent heat fluxes at the various measurement stations of the LITFASS-2003 experiment. A modified Nash-Sutcliffe

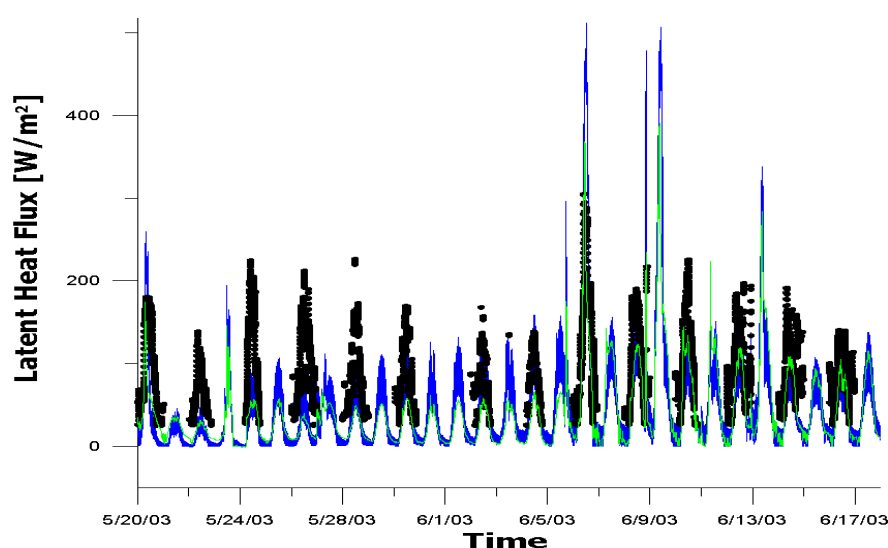


Fig. 6. Mean latent heat flux measurements (verification period) above the LITFASS 2003 area (black points) compared with the LSS TERRA/LM using the standard parameter set (green) and using the MOSCEM-UA-algorithm (pareto set of rank 1; blue).

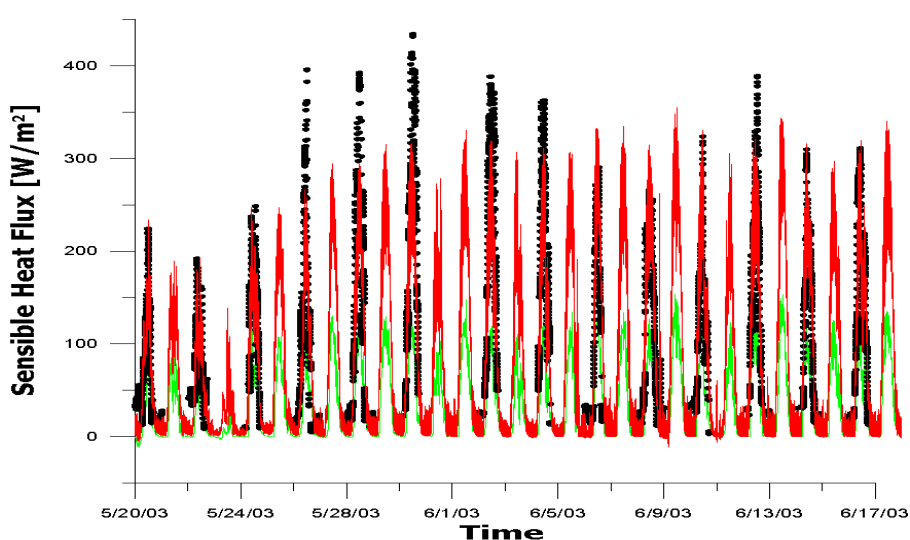


Fig. 7. Mean sensible heat flux measurements (verification period) above the LITFASS 2003 area (black points) compared with the LSS TERRA/LM using the standard parameter set (green) and using the MOSCEM-UA-algorithm (pareto set of rank 1; red).

measure of efficiency, widely used in hydrology for comparing measurements with model results, has been used to rate the flux simulations. In addition to measurements of sensible and latent heat fluxes, the meteorological forcing data were taken at the same sites. The parameter sets found by calibration varied widely and there was no typical classification according to vegetation type. Indeed, parameter sets found by calibration tended to correct for model deficiencies and data errors rather than describe the physical characteristics of the measurement site.

The measured fluxes were combined into a time series of aggregated fluxes by the tile method. Calibration of the land

surface model TERRA/LM with respect to the averaged fluxes gave a range of parameter sets which all allowed TERRA/LM to simulate the area averaged fluxes in much better agreement with observations than the standard parameter set of the operational model.

This study was limited to the rather short time period of the LITFASS-2003 experiment and to one grid box of the operational weather forecast model. Extending this methodology in time and space could lead to an improved description of the surface fluxes in LM and to a more realistic representation of the boundary layer.

## Acknowledgements

Financial support for this study was provided by the Bundesministerium für Bildung und Forschung under grant no. 01 LD 0103. This study is part of the project EVA-GRIPS (Evaporation over Grid and Pixel Scale). Special thanks are presented to all participants in the field experiment LITFASS 2003 and to L.A. Bastidas of the University of Arizona for providing the software packages for the single-criteria (SCE-UA) and multi-criteria calibration (MOSCEM). Two anonymous reviewers are thanked for their helpful remarks.

## References

- Avisar, R. and Pielke, R.A., 1989. A parameterization of heterogeneous land surface for atmospheric numerical models and its impact on regional meteorology. *Mon. Weather Rev.*, **117**, 2113–2136.
- Bastidas, L.A., Gupta, H.V. and Sorooshian, S., 2001. Bounding the parameters of land-surface schemes using observational data. In: *Land surface hydrology, meteorology and climate*, V. Lakshmi, J. Albertson and J. Schaake (Eds.), Water Science and Application 3, American Geophysical Union, 65–76.
- Bastidas, L.A., Gupta, H.V., Sorooshian, S., Shuttleworth, W.J. and Yang, Z.L., 1999. Sensitivity analysis of a land surface scheme using multicriteria methods. *J. Geophys. Res.*, **104**, D16, 19481–19490.
- Bange, J., Wilken, S., Spieß, T. and Zittel, P., 2004. Calibrated Surface Temperature Maps of Heterogeneous Terrain Derived from Helipod and German Air Force Tornado Flights during LITFASS-2003, *Fourth Study Conference on BALTEX, Gudhjem, Denmark*; International BALTEX Secretariat Publication No.29, ISSN 1681-6471. 39–40.
- Beyrich, F., Adam, W.K., Bange, J., Behrens, K., Berger, F.H., Bernhofer, C., Bösenberg, J., Dier, H., Foken, T., Gödecke, M., Görsdorf, U., Güldner, J., Hennemuth, B., Heret, C., Huneke, S., Kohsiek, W., Lammert, A., Lehmann, V., Leiterer, U., Leps, J.-P., Liebenthal, C., Lohse, H., Lüdi, A., Mauder, M., Meijninger, W.M.L., Mengelkamp, H.-T., Queck, R., Richter, S.H., Spieß, T., Stiller, B., Tittbrand, A., Weissensee, U. and Zittel, P., 2004. *Verdunstung über einer heterogenen Landoberfläche – Das LITFASS – 2003 Experiment Ein Bericht, Arbeitsergebnisse Nr. 79*, Deutscher Wetterdienst, Offenbach, Germany. ISSN 1430-0281.
- Beyrich, F., Bange, J., Berger, F.H., Bernhofer, C., Foken, T., Hennemuth, B., Huneke, S., Kohsiek, W., Leps, J.-P., Lohse, H., Lüdi, A., Mauder, M., Meijninger, W. and Mengelkamp, H.-T., 2004b. Energy and water vapor fluxes over heterogeneous land surfaces: The LITFASS-2003 experiment, paper 9.1, *16<sup>th</sup> Symposium on Boundary Layers and Turbulence*, American Meteorological Society, Portland/Maine, USA.
- Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J. and M.F. Wilson, M.F., 1986. *Biophere-Atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model*, NCAR Technical Note NCAR/TN-275+STR, Atmospheric Analysis and Prediction Division, National Center for Atmospheric Research, Boulder, Co., USA.
- Duan, Q., Sorooshian, S. and Gupta, V.K., 1994. Optimal use of the SCE-UA global optimisation method for calibrating watershed models. *J. Hydrol.*, **158**, 265–284.
- DWD, 1995. *Documentation of the EM/DM-System*. Research Department, Deutscher Wetterdienst, Offenbach, Germany.
- Foken, T. and Oncley, S.P., 1995. Workshop on instrumental and methodological problems of land-surface flux measurements. *Bull. Amer. Meteorol. Soc.*, **76**, 1191–1193.
- Foken, T. and Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. *Agr. Forest Meteorol.*, **78**, 83–105.
- Foken, Z., Göckede, M., Mauder, M., Mahrt, L., Amiro, B.D. and J.W. Munger, J.W., 2004. Post-field data quality control. In: *Handbook of Micrometeorology: A Guide for Surface Flux Measurements*, X. Lee, W. Massman and B.E. Law (Eds.) Kluwer, Dordrecht. 181–208.
- Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Co., Reading, MA, USA. 412pp.
- Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1998. Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information. *Water Resour. Res.*, **34**, 751–763.
- Gupta, H.V., Bastidas, L.A., Sorooshian, S., Shuttleworth, W.J. and Yang, Z.L., 1999. Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res.*, **104**, D16, 19491–19503.
- Henderson-Sellers, A., Pitman, A.J., Love, P.K., Irannejad, P. and Chen, T., 1995. The Project for Intercomparison of Land Surface Parametrisation Schemes (PILPS) Phases 2 and 3. *Bull. Amer. Meteorol. Soc.*, **76**, 489–503.
- Houser, P.R., Gupta, H.V., Shuttleworth, W.J. and Famiglietti, J.S., 2001. Multiobjective calibration and sensitivity of a distributed land surface water and energy balance model. *J. Geophys. Res.*, **106**, D24, 33421–33433.
- Louis, J.F., 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Lay. Meteorol.*, **17**, 187–202.
- Mengelkamp, H.-T. and the EVA-GRIPS-Team, 2004. *EVA-GRIPS: Regional Evaporation at Grid and Pixel Scale over Heterogeneous Land Surfaces*. Fourth Study Conference on BALTEX, Gudhjem, Denmark; International BALTEX Secretariat Publication No.29, ISSN 1681-6471. 35–36.
- Noilhan, J. and Planton, S., 1989. A simple parameterization of land surface processes for meteorological models. *Mon. Weather Rev.*, **117**, 536–549.
- Schrodin, R. and E. Heise, E., 2001. *The Multi-Layer Version of the DWD Soil Model TERRA/LM*, Technical Report No.2, Consortium for Small-Scale Modelling (COSMO).
- Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W. and Sorooshian, S., 2003. Effective and efficient algorithm for multiobjective optimisation of hydrologic models. *Water Resour. Res.*, **39**, doi:10.1029/2002WR001746.
- Wilczak, J.M., Oncley, S.P. and Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. *Bound.-Lay. Meteorol.*, **99**, 127–150.
- Wood, E.F., 1995. Scaling behaviour of hydrological fluxes and variables: empirical studies using a hydrological model and remote sensing data. In: *Scale issues in hydrological modelling*, Kalma, Sivaplan (Eds.), Wiley, Chichester, UK. 104pp.
- Zitzler, E. and L. Thiele, L., 1999. Multi-objective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evolut. Comput.*, **3**, 257–271.